

NOVEL GENE DEFECTIVE IN APECED AND ITS USE**Field of the invention**

The present invention relates to a novel gene, a novel protein encoded by said gene, a mutated form of the gene and to diagnostic and therapeutic uses of the gene or a mutated form thereof. More specifically, the present invention relates to a novel gene defective in autoimmune polyendocrinopathy syndrome type I (APS I), also called autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED) (MIM No. 240,300).

10 Background

Autoimmune polyglandular syndrome type I (APS I), also known as autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), is a rare recessively inherited disease (MIM No. 240,300) that is more prevalent among certain isolated populations, such as Finnish, Sardinian and Iranian Jewish populations. The incidence of the disease among the Finns and the Iranian Jews is estimated to be 1:25000 and 1:9000, respectively, whereas only few cases in other parts of the world are found each year.

APECED is one of the two major autoimmune polyendocrinopathy syndromes. The causing factor of APECED has not yet been identified. The syndrome is characterized by lack of tolerance to numerous self-antigens and can therefore be considered as a prototype of organ-specific autoimmune diseases. In APECED, the patient develops chronic mucocutaneous candidiasis soon after birth, and later several organ-specific autoimmune diseases, mainly hypoparathyroidism, Addison's disease, chronic atrophic gastritis with or without pernicious anemia, and in puberty gonadal dysfunction occur [Ahonen P, Clin. Genet. 27 (1985) 535-542]. An accepted criterion for diagnosis of APECED is the presence of at least two of the three main symptoms, Addison's disease, hypoparathyroidism and candidiasis, in patients [Neufeld, M. *et al.*, Medicine 60 (1981) 355-362]. Immunologically, the major findings are the presence of high-titer serum autoantibodies against the effected organs, antibodies against *Candida albicans*, and low or lacking T-cell responses toward candidal antigens [Blizzard, R. M. and Kyle M., J. Clin. Invest. 42 (1963) 1653-1660; Arulanantham, K. *et al.*, New Eng. J. Med. 300 (1979) 164-168; Krohn, K. *et al.*, Lancet 339 (1992) 770-773; Uibo R. *et al.*, J. Clin. Endocrinol. Metab. 78 (1994) 323-328]. The disease usually occurs in

- 100-10309-100

- at present the diagnosis of APECED is based mainly on developed clinical symptoms and typical clinical findings, e.g. the presence of autoantibodies against adrenal cortex or steroidogenic enzymes P450c17 and/or P450scc. The linkage analysis is seldom used. Further, means for natal or presymptomatic diagnosis of the disease are not easily available, since the linkage analysis provides only an indirect data through known gene markers and requires samples from several family members in several generations. Additionally, the linkage analysis is tedious and can be performed only in specialized laboratories by highly-skilled personnel.
- Also the mapping of the carriers of the disease gene is presently based on the linkage analysis and thus not readily available.

Summary of the invention

- We have now identified a novel gene encoding a novel zinc finger protein, designated as autoimmune regulator 1 or AIR-1, which is mutated in APECED. The novel gene and protein allow further development of the diagnosis and therapy of diseases related to immune maturation and regulation of immune response towards self and nonself, such as APECED.

- The object of the invention is to provide means which are useful in a diagnostic method and a gene therapeutic method in the diagnosis and treatment of diseases related to immune maturation and regulation of immune response towards self and nonself, such as APECED.

- Another object of the invention is to provide a novel method for the diagnosis of diseases related to immune maturation and regulation of immune response towards self and nonself, such as APECED, including the pre- and postnatal diagnosis and the mapping of the carriers, the method being easy and reliable to perform.

- The present invention relates to an isolated DNA sequence comprising the sequence id. no. 1 or a functional fragment or variant thereof, or a functionally equivalent isolated DNA sequence hybridizable thereto, the DNA sequence being associated with diseases related to immune maturation and regulation of immune response towards self and nonself, such as APECED. Preferably said isolated DNA sequence includes a gene defect responsible for APECED.

- The present invention also relates to a protein comprising the amino acid sequence id. no. 2 or a functionally equivalent fragment or variant thereof, the protein being associated with diseases related to immune maturation and

regulation of immune response towards self and nonself, such as APECED. Said protein has distinct structural motifs, including the PHD finger motif (PHD), the LXXLL motif (L), proline-rich region (PRR), and cystein-rich region (CRR).

- 5 The present invention further relates to a method for the diagnosis of diseases related to immune maturation and regulation of immune response towards self and nonself, such as APECED, comprising detecting in a biological specimen the presence of a DNA sequence comprising the sequence id. no. 1 or a functional fragment or variant thereof, or a functionally
10 equivalent DNA-sequence hybridizable thereto, the DNA sequence being associated with diseases related to immune maturation and regulation of immune response towards self and nonself, such as APECED.

- The present invention further relates to the use of the above-identified DNA-sequences in the diagnosis of diseases related to immune
15 maturation and regulation of immune response towards self and nonself, such as APECED.

- The present invention further relates to a method for the diagnosis of diseases related to immune maturation and regulation of immune response towards self and nonself, such as APECED, comprising detecting in a
20 biological specimen the presence or the absence of a protein comprising the sequence id. no. 2 or a functionally equivalent fragment thereof, the protein being associated with diseases related to immune maturation and regulation of immune response towards self and nonself, such as APECED.

- The present invention further relates to the use of the above-identified protein or a functionally equivalent fragment thereof in the diagnosis
25 of diseases related to immune maturation and regulation of immune response towards self and nonself, such as APECED.

- The present invention further relates to the use of the above-identified DNA sequences in gene therapy or for the preparation of a
30 pharmaceutical preparation useful in a gene therapy method of diseases related to immune maturation and regulation of immune response towards self and nonself, such as APECED.

Brief description of the drawings

- Figure 1 shows a physical map of the APECED gene locus in the
35 chromosome 21q22.3. Cosmids D1G8, D40G11, D9G11, D28B11, and D4G11, overlapping clones used for the genomic sequencing [Kudoh, J. *et al.*,

DNA Res. 4 (1997) 45 -52] are indicated by horizontal lines. The APECED gene located just proximal to the 5' end of the neighboring gene PFKL is indicated by a solid arrow. N indicates *NotI* sites. DNA marker D21S1912 is shown as open box.

Figure 2 shows the structures of the APECED gene and AIR proteins. (A) Cloning strategy of APECED cDNAs and the order of the exons in the APECED gene. DNA fragments amplified by PCR and 3'- and 5'-RACE are indicated by the lines. Exon 1' is the 5'-noncoding exon of the AIR-2 and AIR-3. An additional alternative splicing of AIR-3 in exon 10, resulting in an amino acid change in its downstream, is indicated by vertical lines. Each exon, except exon 1', is bordered by the common splice site consensus sequence, ag:gt. Mutations in the exon 2 and exon 6 are indicated by the arrows. (B) Schematic presentation of the three AIR proteins showing distinct structural motifs, including the PHD finger motif (PHD), the LXXLL motif (L), proline-rich region (PRR), and cystein-rich region (CRR).

Figure 3 shows electropherograms showing the sequence surrounding the mutations in the APECED gene. (A) Mutation analysis of a Swiss APECED family. The parents are heterozygous for the allele (normal "C" and abnormal "T"). The affected boy and girl show the "C" to "T" transition resulting in the "Arg" to "Stop" nonsense mutation at amino acid position 257. (B) Mutation analysis of two Finnish APECED patients. The patient MP is homozygous for the mutant allele (left), NP is heterozygous for the allele (right). (C) The patient NP shows the "A" to "G" transversion resulting in the "Lys" to "Glu" missense mutation at amino acid position 42. FLEB is a normal control.

Figure 4 shows the result of a restriction enzyme *TaqI* digestion assay demonstrating the R257stop mutation. Four APECED patients [HP1 (lane 1), HP2 (lane 2), NP (lane 6), and MP (lane 8)], the mothers of two families [HM (lane 5) and NM (lane 7)], two healthy siblings [HN1 (lane 3) and HN2 (lane 4)] of family H and normal controls [C1, C2 and C3 (lanes 9-11)] are shown. The APECED patients HP1, HP2 and MP are homozygotes for the R257stop mutation. The APECED patient NP is heterozygous for the R257stop mutation but is carrying a mutation at a different position in another allele of the APECED gene (shown above in Fig. 3C). Both mothers (HM and NM) and two healthy siblings (HN1 and HN2) are heterozygous for the R257stop mutation and therefore carriers of APECED but are not having the

Figure 5 shows an amino acid sequence alignment for the PHD finger motif of AIR-1, Mi-2, and TIF1. The consensus amino acid residues conserved in the PHD finger motif is indicated by the bold letters underneath. The residues that are identical with AIR-1 (aa 299-340) are shown by the dots. GenBank accession nos. of Mi-2 and TIF1 are X86691 and AF009353, respectively.

Figure 6 is a Western blot showing the expression of AIR-1 in fetal liver. A sample of fetal liver was run on PAGE, transferred to a nitrocellulose filter and probed with sera as follows: Lane 1, control mouse serum, lane 2, control mouse serum absorbed with peptide AIR-1/2 (sequence id. no. 25), lanes 3 and 4, serum from a mouse immunized with peptide AIR-1/2 for four and six weeks, respectively and absorbed with peptide AIR-1/2, lanes 5 and 6, unabsorbed serum from a mouse immunized with peptide AIR-1/2 for four and six weeks, respectively. The strong band seen in lanes 5 and 6 represent the AIR-1 protein with a molecular weight of approx. 58 kD, the lower band is an approx. 20 kD breakdown product of the AIR protein. The bands seen in all lanes are non-specific.

Figure 7 shows the expression of the APECED mRNA (7A) or the AIR protein (7B, 7C and 7D) demonstrated by in situ hybridization (7A) or by immunohistochemistry (7B, 7C and 7D). Figure 7A shows APECED mRNA positive cells scattered in the medullary region of human thymus. Figure 7B shows similar cells with the same localization now stained for the AIR protein. Figure 7C is a higher magnification of 7B, showing the localization of the AIR protein in the nuclei. Note the speckled localiation pattern in the nuclei. Figure 7D shows the cytoplasmic localization of the AIR protein in a few cells in lymph node medulla.

node medulla.

30 Figure 8 shows the phenotypic characterization of the APECED reactive cells in thymus by double-immunofluorescence. The AIR protein is seen as red colour in the nuclei, forming typical speckled pattern with nuclear dots. In Figure 8A, the co-staining is with an antibody recognizing low molecular weight markers (AE1). The APECED positive cells fall into two

35 types, one is expressing cytokeratin and is thus epithelial cell, the other one is non-epithelial and do not co-express cytokeratins. In Figure 8B an APECED

positive cell co-expresses a marker (CD83) typical for cells belonging to monocyte-macrophage-dendritic cell lineage.

Figure 9 shows the expression of the AIR protein, demonstrated by immunofluorescence, in mature, activated dendritic cells from peripheral blood.

- 5 The expression of the AIR protein shows as distinct dots in the nuclei of dendritic cells.

Detailed description of the invention

The present invention is based on studies aiming for the identification and characterization of the gene defect in APECED. In the sequence studies, a cosmid/BAC (bacterial artificial chromosome) contig of 520 kb covering four gene markers D21S1460-D21S1912-PFKL-D21S154 [Kudoh, J. *et al.*, DNA Res. 4 (1997) 45-52] was constructed, and genomic sequencing in this region was performed [Kawasaki, K. *et al.*, Genome Res. 7 (1997) 250-261]. From this genomic sequence information the distance between D21S1912 and PFKL was determined to be approximately 140 kb (Fig. 1).

1). Using a computer program, such as GRAIL and GENSCAN [Uberbacher, E. C. and Mural, R. J., Proc. Natl. Acad. Sci. USA 88 (1991) 11261-11265; Burge, C. and Karlin, S., J. Mol. Biol. 268 (1997) 78-94], gene
20 screening in the partial sequencing data within this region was performed. GENSCAN predicted several genes between D21S1912 and PFKL. One of these genes located just proximal to the PFKL gene contained the previously trapped exon HC21EXc33 [Kudoh, J. *et al.*, DNA Res. 4 (1997) 45-52] or MDC04M06 [Chen, H. *et al.*, Genome Res. 6 (1996) 747-760]. A set of primers
25 for polymerase chain reaction (PCR) was then designed from the predicted exons. The PCR screening of various cDNA libraries using these primers allowed the isolation of a cDNA clone containing the exon HC21EXc33 (exon 13) from the thymus cDNA library (Fig. 2A).

A 3'-rapid amplification of cDNA ends (3'-RACE) and 5'-RACE using MarathonTM cDNA Amplification Kit (Clontech Laboratories Inc, California, USA) according to manufacturer's protocol from the thymus cDNA library was performed using a primer c33F (sequence id. no. 7) and a primer 1R (sequence id. no. 8), respectively.

Sequencing analysis revealed a unique sequence of 2027 bp in
35 overlapping PCR products that contains a 1635-bp open reading frame (ORF)
from methionine at nt 128 to a TAG stop codon at nt 1763 encoding a predic-

ted novel protein designated AIR-1, for autoimmune regulator 1. AIR-1 encodes a protein of 545 amino acids with a predicted isoelectric point of 7.32 and a calculated molecular mass of 57,723 (Fig. 2B).

A 5'-RACE from the thymus cDNA using a primer 4R (sequence id. no. 9) resulted in an alternatively spliced product. Furthermore, two types of the cDNA clones were amplified with a primer pair 3F/c33R (sequence id. no. 10/sequence id. no. 11) and these clones encode for AIR-2 and AIR-3 proteins, sequence id. no. 4 and sequence id. no. 6, respectively (Fig. 2A) (sequence id. no. 3 and sequence id. no. 5). The AIR-2 and AIR-3 proteins consist of 348 and 254 amino acids, respectively (Fig. 2B). These results suggest that the APECED gene is transcribed as at least three types of mRNA by alternative splicing and/or use of an alternative 5' exon within the gene. RT-PCR analysis [Griffin, H. G. and Griffin, A. M., PCR Technology. Current Innovations, CRC Press, 1994] revealed that the AIR-1 transcript is also expressed in fetal liver (data not shown).

The APECED gene is approximately 13-kb in length and contains 15 exons, including the exon 1' specific to AIR-2 and AIR-3. It is transcribed in the direction of centromere to telomere (Figs 1, 2A). Based on this information, PCR primers were designed to amplify each exon from the genomic DNA and a mutation analysis of Swiss and Finnish APECED families was performed. Sequence comparison identified two mutations in the APECED gene of the patients (Fig. 3). The first mutation changes an Arg codon (CGA) to a stop codon (TGA) at amino acid position 257 in exon 6. This mutation was designated as R257stop mutation. The second mutation is a missense mutation that derived from the maternal chromosome in one Finnish patient (NP): a Lys codon (AAG) changes to a Glu codon (GAG) at amino acid position 42 in exon 2. This mutation is designated as K42E mutation (Figs 2A, 3C).

The R257stop mutation destroys a *TaqI* restriction enzyme site and the K42E mutation introduces a novel *TaqI* site. Thus these two mutations can be easily demonstrated in one or both alleles by *TaqI* digestion or by digestion using another enzyme cleaving at the recognition site 5'-TCGA-3' (Fig. 4).

The AIR-1 protein has strong homology in certain domains to the major autoantigens (Mi-2) associated with the autoimmune disease dermatomyositis [Seeig, H. P. *et al.*, Arthritis Rheum. 38 (1995) 1389-1399; Ge, Q. *et al.*, J. Clin. Invest. 96 (1995) 1730-1737], Sp140, a protein from the nuclear body, an organelle involved in the pathogenesis of certain types of

leukemia, and which is also the target of antibodies in the serum of patients with the autoimmune disease primary biliary cirrhosis [Bloch, D. B. *et al.*, J. Biol. Chem. 271 (1996) 29198-29204]. In addition, the homologies extend to other nuclear proteins such as TIF1 [Le Douarin, B. *et al.*, EMBO J. 14 (1995) 2020-2033], LYSP100 [Dent, A. L. *et al.*, Blood 88 (1996) 1423-1426], and putative yeast and *C. elegans* proteins. The AIR-1 protein homologies are principally in two PHD finger motifs (amino acid 299 to 340 and 434 to 475) (Fig. 5). AIR-1 also contains a proline-rich region (amino acid 350 to 430) (Fig. 2B). The PHD finger is a cysteine-rich structure that is distinguished from the RING finger (C3HC4) and LIM domain (C2HC5) because it contains a consensus of C4HC3. [Aasland, R. *et al.*, Trends Biochem. Sci. 20 (1995) 56-59]. The PHD finger motif is found in a number of chromatin-associated proteins such as HRX that is involved in the t(11:17) translocation in acute leukemia [Chaplin, T. *et al.*, Blood 86 (1995) 2073-2076]. The proline-rich region is assumed to be involved in protein-protein interaction or DNA binding. The presence of the PHD finger and proline-rich regions indicates a function for AIRs as transcription regulatory proteins. However, the AIR proteins have no apparent nuclear translocation signal, and thus other proteins containing such signal may interact with AIR to translocate it to the nucleus. In fact, the AIR proteins also have the LXXLL motif that is a signature sequence to bind to nuclear receptors [Heery, D. M. *et al.*, Nature 387 (1997) 733-736] (Fig. 2B).

The clinical picture of APECED and the observed immunological abnormality with strong autoimmune response towards several target organs and antigens suggest that the product of the APECED gene has a central role in immune (ontogeny) maturation and regulation of immune response towards self and nonself.

According to the diagnostic method of the invention, the presence of the defective APECED gene can be detected from a biological sample by any known detection method suitable for detecting mutations. Such methods include the method described by Saiki *et al.* [Proc. Natl. Acad. Sci USA 86 (1989) 6230-6234] utilizing hybridization to an allele specific oligonucleotide probe, or modifications thereof; the method described by Newton, C. R. *et al.* [Nucl. Acids Res. 17 (1989) 2503-2516] using the DNA sequences or DNA-fragments of the invention as probes; the solid phase minisequencing method described by Syvänen *et al.* [Genomics 8 (1990) 684-692] in which use is made of a biotinylated probe; or the oligonucleotide ligation method described

by Landegren, U. *et al.* [Science 241 (1988) 1077-1080]. Methods include the denaturing gradient gel electrophoresis (DGGE) [Fischer, S.G. and Lerman, L.S., PNAS 80 (1983) 1579-1583] or a modification of this method, constant denaturant gel electrophoresis (CDGE) [Hoving *et al.*, Genes Chromosomes
5 Cancer 5 (1992) 97-103]. The mutation separation principle of DGGE and CDGE is based on the melting behavior of the DNA double helix of a given fragment.

Since the mutations of the APECED gene involve a site sensitive to *TaqI* digestion, the mutation are preferably detected in one or both alleles by
10 *TaqI* digestion or by digestion using another enzyme cleaving at recognition site 5'-TCGA-3' The chemical mismatch cleavage for mutation analysis can be used [Grompe, M. *et al.*, Proc. Natl. Acad. Sci. USA 86(15)(1989) 5888-5892].

In the diagnostic method of the invention the biological sample can be any tissue or body fluid containing cells, such as blood, e.g. umbilical cord
15 blood, separated blood cells, such as lymphocytes, B-cells, T-cells etc., biopsy material, such as fetal liver or thymus biopsy, sperm, saliva, etc. The biological sample can be, where necessary, pretreated in a suitable manner known to those skilled in the art.

When the DNA sequence of the present invention is used
20 therapeutically any techniques presently available for gene therapy can be employed. Accordingly, in the technique known as *ex vivo* therapy patient cells (e.g. umbilical cord blood from the fetus) with the defective gene are taken from the patient, DNA sequences encoding the normal (healthy) gene product incorporated in a carrier vector are transduced or transfected to the cells and
25 the cells are returned to the patient. If the techniques known as *in situ* therapy is used, the DNA sequences encoding the normal gene product are first inserted to a suitable carrier vector, and the carrier is then introduced to the affected tissue, such as peripheral blood, liver or bone marrow. The carrier vector used can be a retrovirus vector, an adeno virus vector, an adeno
30 associated virus (AAV) vector or an eucaryotic vector. The therapy can be performed intra utero or during adult life. Depending on the cells to be treated these techniques lead either to a transient cure, where cells from affected organ are treated, or to a permanent cure, in case of the treatment of stem cells.

35 The present invention provides means for an easy and more rapid diagnosis of the diseases related to immune maturation and regulation of

immune response towards self and nonself, such as APECED, and, specifically, enables prenatal diagnosis and carrier diagnosis. Furthermore, it provides a background for therapy.

The invention is now elucidated by the following non-limiting
5 examples.

Example 1

Localization of the APECED gene

Genomic sequencing of cosmid DNAs was performed by the shotgun method described by Kawasaki, K. *et al.*, Genome Res. 7 (1997) 250-
10 261. Cosmids D1G8, D40G11, D9G11, D28B11, and D4G11 and gene marker D21S1912 are described by Kudoh, J. *et al.*, DNA Res. 4 (1997) 45-52].

cDNA cloning

The phage DNAs prepared from human thymus cDNA library (Clontech, HL1127a) were used as a PCR template. 20 ng of phage DNA
15 which represents approximately 4×10^8 phages was added to a 10 ml of reaction mixture containing 1x buffer [16mM $(\text{NH}_4)_2\text{SO}_4$, 50mM Tris-HCl, pH 9.2, 1.75 mM MgCl_2 , 0.001% (w/v) gelatin), 0.2 mM each of dNTPs, 1M Betaine (Sigma), 0.35 U of Tap and Pwo DNA polymerase (EXpand Long Template PCR System, Boehringer Mannheim), and 0.5 mM of each of the
20 primers, 2F and c33R, 2F and 4R, and 2F' and 2R', respectively.

The cDNA fragment was amplified by PCR using the following conditions: 94°C for 3 min., 35 cycles of 94°C for 30 sec, 60 °C for 30 sec in 2F/c33R and 2F/4R or 65°C for 30 sec in 2F'/2R', and 68°C for 90 sec. 3'- and 5'-RACE were carried out by Marathon cDNA Amplification Kit (Human
25 Thymus; Clontech). PCR reaction was performed in a 10 µl volume containing 1x buffer (50 mM KCl, 10 mM Tris-HCl, pH 8.3, 1.5 mM MgCl_2 , 0.001% (w/v) gelatin), 0.2 mM each of dNTPs, 0.25 U of AmpliTaq Gold polymerase (Perkin-Elmer), and 0.5 mM of each of the exon-specific primers. 3'-RACE product was amplified by PCR with the following conditions: 95°C for 9 min., 35 cycles
30 of 94°C for 30 sec, 60°C for 30 sec, and 72°C for 30 sec.

The cDNA fragments were sequenced by the dye deoxy terminator cycle sequencing method (according to ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction Kit protocol P/N 402078, Perkin Elmer Corporation, California) using specific primers, 2F and c33R, and AmpliTaq/FS

095086553 140300

DNA polymerase (Perkin-Elmer), and then analyzed by using an automatic DNA sequencer (Applied Biosystems 377). Primer sequences used were

- 1R: 5'-GTTCCCGAGTGAAGGCGCTGC-3' (sequence id. no. 8)
 2F: 5'-GGATTCAGACCATGTCAGCTTCA-3' (sequence id. no. 12)
 3F: 5'-GAGTTCAGGTACCCAGAGATGCTG-3' (sequence id. no. 10)
 c33R: 5'-CTCGCTCAGAAGGGACTCCA-3' (sequence id. no. 11)
 4R: 5'-AGGGGACAGGCAGGCCAGGT-3' (sequence id. no. 9)
 2F': 5'-GTGCTGTTCAAGGACTACAAC-3' (sequence id. no. 13)
 2R': 5'-TGGATGAGGATCCCCTCCACG-3' (sequence id. no. 14)
 AP1: 5'-CCATCCTAATACGACTCACTATAGGGC-3' (sequence id. no. 15) and
 c33F: 5'-GATGACACTGCCAGTCACGA-3' (sequence id. no. 7).

Example 2

15 Mutation analysis of the APECED gene

For the mutation analysis the DNA samples were purified from peripheral blood mononuclear cells from patients with APECED and from suspected carriers of APECED and from normal healthy controls (according to Sambrook *et al.* 1989, Molecular Cloning. A Laboratory Manual. CSH Press) and subjected to PCR using primers specific for all identified exons.

For sequencing the mutated exons, PCR fragments, 6F/6R in exon 6 and 49300F/49622R in exon 2, were amplified by PCR with the following conditions: 95°C for 9 min., 35 cycles of 94°C for 30 sec, 60°C for 30 sec and 72°C for 30 sec, and 94°C for 3 min., 35 cycles of 94°C for 30 sec, 60°C for 30 sec, and 68°C for 30 sec, respectively. The PCR products were sequenced using specific primers

- 6F: 5'-TGCAGGCTGTGGGAACTCCA-3' (sequence id. no. 16)
 6R: 5'-AGAAAAAGAGCTGTACCCTGTG-3' (sequence id. no. 17)
 3R: 5'-TGCAAGGAAGAGGGGCGTCAGC-3' (sequence id. no. 18)
 49300F: 5'-TCCACCACAAGCCGAGGAGAT-3' (sequence id. no. 19) and 49622R: 5'-ACGGGCTCCTCAAACACCACT-3' (sequence id. no. 20).

In the mutation analysis by sequencing, two Swiss and three Finnish (HP1, HP2 and MP) patients with APECED were homozygous for R257stop allele, whereas one Finnish patient (NP) was heterozygous for this mutation (Fig. 3A, B). The R257stop mutation of NP was derived from the

paternal chromosome. The second mutation, K42E mutation, was found in one Finnish patient (NP): a Lys codon (AAG) changes to a Glu codon (GAG) at amino acid position 42 in exon 2.(Figs 2A, 3C). This mutation derived from the maternal chromosome.

5 Example 3

Restriction enzyme *TaqI* analysis of two mutations in exons 2 and 6 of APECED gene

Analysis of the mutation sites in exons 2 and 6 in large series of individuals was performed using the restriction enzyme *TaqI*. The *TaqI* digestion for exons 2 and 6 was done as follows. Ten microlitres of
10 digestion for exons 2 and 6 was done as follows. Ten microlitres of amplification product was incubated at 65°C for 1 hour in 20 µl of reaction mixture containing 1x *TaqI* digestion buffer (New England Biolabs, NY, 100 µl/ml of BSA and 10U of *TaqI* enzyme (New England Biolabs, NY). After the digestion fragments were separated in 1.5% agarose gel and visualized by
15 EtBr staining.

For exon 2, the fragment containing the mutation site K42E was amplified with primers GR1/2F and GR1/2R with the following conditions: 95°C for 3 min., 35 cycles of 94°C for 30 sec, 62°C for 30 sec and 72°C for 1 min. The 1x reaction mix used contained 50 mM KCl, 10 mM Tris-HCl, pH 8.3, 1.5
20 mM MgCl₂, 0.001% (w/v) gelatin), 0.2 mM each of dNTPs, 0.25 U of Dynazyme (Finnzymes, Finland), and 0.5 mM of each of the exon-specific primers. The normal allele produces a 312 bp fragment whereas the mutated allele gives a 133 bp and a 179 bp fragment. Primer sequences for GR1/2F and GR1/2R are 5'-TGGAGATGGGCAGGCCGCAGGGTG (sequence id. no.
25 21) and 5'-CAGTCCAGCTGGGCTGAGCAGGTG (sequence id. no. 22), respectively.

For exon 6, the fragment containing the R257stop mutation site was amplified with primers GR1/5IF and GR1/5IR with the same conditions described for exon 2 (see above). The normal allele produces a 225 bp frag-
30 ment whereas the mutated allele gives a 285 bp fragment. Primer sequences for GR1/5IF and GR1/5IR are 5'-GCGGCTCCAAGAAGTGCATCCAGG (sequence id. no. 23) and 5'-CTCCACCCTGCAAGGAAGAGGGGC (sequence id. no. 24), respectively.

The screening of 50 Finnish and 50 Swiss healthy individuals did
35 not reveal R257stop or K42E mutations by *TaqI* digestion. Similarly, PCR ana-

lysis of 20 unaffected Japanese was performed and no mutations were found in any of these positions. These results demonstrate that the APECED gene is responsible for the pathogenesis of APECED.

Mutations were found in the AIR-1 transcript but not in the AIR-2 and AIR-3 transcripts from all the APECED patients tested. Two Swiss and three Finnish (HP1, HP2 and MP) patients who are homozygous for the R257stop mutation completely lack functional AIR-1 protein but still have intact AIR-2 and AIR-3 proteins.

One common mutation seems responsible for the genetic defect in approximately 90% of the Finnish APECED cases and a haplotype analysis with the markers D21S141, D21S1912 and PFKL shows that the R257stop mutation is likely to be this common mutation [Björnses, P. *et al.*, Am. J. Hum. Genet. 59 (1996) 879-886].

Example 4

15 Analysis of the AIR protein expression

In this example, synthetic peptides representing amino-acid sequences of the AIR-1 protein, were used to generate a polyvalent mouse antiserum against the AIR-1 protein.

For the peptide synthesis, two peptides were chosen according to the antigenicity prediction by Pepsort program (GCC package, Wisconsin, USA). The peptides AIR-1/2 and AIR-1/6 (TLHLKEKEGCPQAFH, sequence id. no. 25 and GKNKARSSSGPKPLV, sequence id. no. 26, respectively) representing exons 2 and 6, respectively, of the APECED gene were synthesized onto a branched lysine core (Fmoc8-Lys4-Lys2-Lys-betaAla-Wang resin, Calbiochem-Novabiochem, La Jolla, Ca, USA) resulting in an octameric multiple antigen peptide (MAP) [Tam, J. P. *et al.*, Proc. Natl. Acad. Sci. USA 85 (1988) 5409-5413; Adermann, K. *et al.*, in Solid Phase Synthesis, Biological and Biomedical Applications, pp. 429-432, Ed. R. Epton, Mayflower Worldwide Ltd., Birmingham, 1994], Syntheses were performed by Fmoc (N-(9-fluorenyl)methoxycarbonyl) chemistry on a simultaneous multiple peptide synthesizer (SMPS 350, Zinsser Analytic, Frankfurt, Germany). Purity of MAPs was analyzed by reverse-phase HPLC (System Gold, Beckman Instruments Inc, Fullerton, CA, USA).

To obtain murine polyclonal antibodies, eight-week old Balb/c mice were immunized with an intraperitoneal injection of 25 micrograms of each

peptide in 0,4 ml of a 1:1 mixture of Freund's Complete Adjuvant (Difco Laboratories, Detroit, MI, USA) and physiological saline (NaCl, 0,15 M). One month later the animals were boosted with an intramuscular injection of 35 micrograms of antigens in Freund's incomplete adjuvant and saline (1:1) (0,2 ml were distributed into four sites). Three weeks later the peptides in a dose of 50 micrograms/mouse were administered intravenously and sera were obtained 7 days later.

For the production of EBV transformed B-cells, peripheral blood leukocytes were obtained from healthy control persons. The B-cells were transformed with EBV (Epstein-Barr virus) using standard protocol, and the cell lines were maintained in RPMI 1640, supplemented with 10% FCS (fetal calf serum). An aliquot of cells were stimulated for 12 hours with 10 mg/ml of phytohemagglutinin (PHA) to obtain mitogen-activated T-cells.

Tissue samples were obtained from stillborn fetuses at six months gestational age. Fetal liver, spleen, thymus and lymphnodes were homogenized, the homogenates were cleared with centrifugations (20 000 rpm for 20 minutes) and the samples were used for western blot analysis.

For analysis of polyclonal sera, Elisa and western blot analysis were performed. Microtitre ELISA plates (Maxisorp, Nunc, Roskilde, Denmark) were coated with the peptides (1 micrograms /well in PBS, pH 7,5) at 4°C overnight and blocked with 2 % of BSA in PBS. The plates were then incubated with titrated mouse immune sera and normal (control) sera at room temperature for 4 h. Finally the bound peptide-specific antibodies were detected by use of anti-mouse HRP-labelled immunoglobulins (Dako A/S, Denmark) essentially as previously described [Ovod, V. A. *et al.*, AIDS 6 (1992) 25-34].

For western blotting, tissue homogenates, EBV transformed B-cells or PHA-activated T-cells were boiled for 10 minutes in 2x sample buffer (for tissue homogenates: 100 microliters of homogenate mixed with 100 microliters of sample buffer; for cells: one million cells/100 ml of buffer) and analyzed in western blotting as described in Ovod, V. A. *et al.*, *supra*.

The antisera so produced reacted with the AIR-1-protein low amount in normal fetal spleen, thymus and lymphnode as well as, in EBV-transformed B-cells and in PHA-activated T-cells. In the ELISA assay towards the immunogenic peptides, all four mice gave a strong reactivity towards the peptide used for the immunization. In the western blotting analysis using either

For the preparation of antibodies to the AIR protein, the APECED cDNA (sequences 137 - 1774 of sequence id. no. 1) containing a full-coding region was amplified from Marathon human thymus cDNA (Clontech) with primers ExF and ExR2. The primer sequences for ExF and ExR2 were 5'-CCA CCC CAT GGC GAC GGA CG-3' (sequence id. no. 33) and 5'-GGA ATT CGG AGG GGA AGG GGG CCG CCG GA-3' (sequence id. no. 34). The amplified cDNA was digested with NcoI and EcoRI and cloned (pHPAIRE) into pET32a vector (Novagen). The protein was expressed in *E. coli* and purified by His-tag as described by manufacturer (QiaExpress Type IV Kit, Cat No 32149, Qiagen, USA).

To obtain murine polyclonal antibodies, Balb/c mice were immunised essentially as described in Example 4 using 100 micrograms of the bacterially expressed AIR protein with booster doses of 25 and 25 micrograms.

Japanese white rabbits were immunised with a synthetic peptide representing amino acids 526-545 (DGILQWAIQSMARPAAPFPS, sequence id. no. 36) of sequence id. no. 2. The specificities of the antisera were checked with ELISA and Western blotting using standard procedures.

For immunocytochemistry, frozen sections of tissue samples were fixed for 20 min in 4% paraformaldehyde. The AIR antibody (rabbit or mouse) in an appropriate dilution was incubated for 30 min at 37°C, with a biotin conjugated anti-mouse or anti-rabbit secondary antibody (Vector, CA, USA). The biotinylated antibody was revealed by incubating with Texas Red-avidin (Vector, CA, USA) for 30 min at 37°C.

With in situ hybridization, a positive signal was seen in a few cells in thymus medulla (Fig. 7A). The APECED *in situ* -positive cells were infrequent and scattered as single cells in the medulla, but occasionally one or two APECED-expressing cells were seen adjacent to or buried into the Hassal's corpuscles that represent conglomerates of medullary epithelial cells. In the positive cells, APECED mRNA was predominantly localized in the cell nucleus. In human adult lymph node tissues, infrequent cells expressed APECED mRNA in the medulla and occasionally in the paracortical region, too (Fig. 7B). No hybridization signal was seen in the germinal centers.

Immunohistochemistry with mouse and rabbit polyclonal antisera to the AIR protein showed strong reactivity with selected cells in thymus medulla, lymph nodes and fetal liver (Fig. 7C and 7D). The comparison of the reaction

pattern obtained by immunohistochemistry to that obtained by in situ hybridization clearly established that specific, rare cells in thymus medulla and lymph node medulla and paracortex express APECED mRNA and the AIR protein. By either method, neither mRNA nor protein was detected in other adult tissues studied, including the target organs for tissue destruction in APECED (adrenal glands, parathyroid glands, gonads). In human fetal tissues, APECED positive cells were seen, although extremely infrequently, in the stroma of placental chorionic villi and in the sinusoidal area of the liver. In the fetal liver, the APECED positive cells were often localized pairwise like mirror images, suggesting that the cells were undergoing mitosis. Rare APECED expressing cells were also found in fetal thymus but the expression was not observed in other fetal tissues.

At the subcellular level, the AIR protein localized in small nuclear dots in the adult thymus, giving a characteristic speckled pattern (Fig. 7C; and Fig. 8A and 8B), but localized in the cytoplasm of cells in lymph nodes. In the rare positive cells in fetal liver, many of which were mitotic, the AIR protein was localized in the cytoplasm.

Example 6

Characterization of the phenotype of the APECED positive cells in thymus

Double staining with two antibodies was used to further characterize the cell type expressing APECED gene. In view of the fact that dendritic cells (DC) and thymus epithelium are both involved in the regulation of immune maturation, expression of markers for these cells were studied.

For double immunofluorescence detection the AIR staining was performed as described in example 5 with rabbit anti-AIR serum. The slides were then incubated with a second primary antibody [AE1 (Neomarkers, CA, USA), AE3 (Neomarkers, CA, USA), CD11c (Immunotech, France), or CD83 (Immunotech, France)] in an appropriate dilution for 30 min at 37°C, and the reaction was revealed by incubating with a FITC conjugated secondary anti-mouse antibody (Vector, CA, USA) for 30 min at 37°C.

Antibodies reacting with low molecular weight basic (AE1) or high molecular weight acidic (AE3) cytokeratins stained the thymus in a reticular fashion, and the APECED positive cells were seen either buried into this net or in close apposition with the keratin-positive cells. Confocal microscopy

clearly demonstrated that some of the APECED positive cells were cytokeratin positive while some remained negative (Figure 8A). A co-localization was stronger with AE1 than with AE3. The distribution of epithelial (AE1 positive) and non-epithelial APECED expressing cells varied but in most thymus preparates more than half were epithelial.

Less than half of the APECED expressing cells in thymus stained with markers CD11c and CD83 that react with cells of the monocyte-macrophage-dendritic cell lineage. In most cases, the staining reaction was weak but a few cells showed an intensive staining with the given marker (Fig. 8B). CD83 costained 5 to 40 % of the APECED positive cells. Antibody CD11c, reported to be specific for mature dendritic cells, reacted with up to 5 - 10 % of the APECED positive cells. All APECED positive cells were strongly positive for HLA-DR staining, however (data not shown).

These results suggest that in thymus the APECED gene is in fact expressed in two distinct cell populations, one epithelial and the other non-epithelial. The latter cell type is likely the one also expressing the APECED gene in extrathymic lymphoid tissues.

Example 7

APECED expression in stimulated dendritic cells *in vitro*

To show an APECED expression in dendritic cells derived from peripheral blood monocytes that are DC precursors, these cells were cultured at the presence of cytokines using conditions that are known to lead to the expansion and maturation of dendritic cells.

Peripheral blood mononuclear cells were isolated by Ficoll-Hypaque centrifugation, and adherent cells were separated and cultured in the presence of human recombinant GM-CSF (1000 units/ml) and rhIL-4 (1000 units/ml, both from R&D Systems), as described [Schuler, G. and Romani, N., Adv. Exp. Med. Biol. 417 (1997) 7 - 13]. Cells were further cultured for three days with 1/4 V/V of macrophage conditioned media. Cells were harvested at two days intervals and samples were prepared for RT-PCR. For RT-PCR total RNA was purified from DCs by using a commercial kit from Clontech (USA) (Nucleospin RNA Kit) according to manufacturer's instructions. An aliquot of RNA was transferred into cDNA with a commercial kit from Pharmacia (Sweden) (First-strand Synthesis Kit) and PCR for this DNA sample was performed. For PCR the fragment was amplified with primers 5'- GAT CCT

GCT CAG GAG ACG TGA CCC-3' (seq. id. no. 31; 1554 -1577 of seq. id. no. 1) and 5'-GGA CTG AGG AAG GAG GTG TCC TTC -3' (seq. id. no. 35; 1818-1841 of seq. id. no. 1) with the following conditions: 35 cycles of 95°C for 1 min., 62°C for 30 sec and 72°C for 1 min. The 1x reaction mix contained
5 50mM KCl, 10mM Tris-HCl, pH8.3, 1.5mM MgCl₂, 0.001% (w/v) gelatin, 0.2mM each of dNTPs, 0.25 U of Dynazyme (Finnzymes, Finland). A fragment of 287bp was detected by 1.5% agarose electrophoresis.

Cytospin preparations were further made for immunohistochemistry.

During this 7 to 10 days culture period approximately half of the
10 cells developed the characteristic veiled morphology of DC and their phenotypic cell markers (CD11c and CD83) corresponded to mature DCs (Figure 9). The APECED expression was studied by RT-PCR and immunocytochemistry at two to three days intervals. In the starting material, i.e. the adherent cell pool from peripheral blood, no APECED expression was
15 found. After seven days of culture in the presence of GM-CSF and IL-4, RT-PCR showed APECED mRNA expression and immunofluorescence showed a few AIR specific nuclear dots. After an additional 3-day-culture with conditioned medium from macrophage cultures a strong speckled pattern of nuclear AIR expression was seen (Figure 9A). The RT-PCR analysis of the
20 mature (10 days) culture confirmed the AIR protein expression.